WAVELET BASED ECG ARRHYTHMIA CLASSIFICATION

USING GA-SVM METHOD

RAORANE SHWETA SURESH & JADHAV SHARAD P

[1]*Department of Instrumentation, Konkan Gyanpeeth College of Engineering, Karjat, Maharashtra, India*

[2]*Lecturer & HOD, Department of Instrumentation, Ramrao Adik Institute of Technology, Maharashtra, India*

ABSTRACT

This project work presents approach for cardiac arrhythmia disease classification. The GA-SVM method combines both Support Vector Machine (SVM) and Genetic Algorithm approaches. First, twenty two features from electrocardiogram signal are extracted. ECG feature extraction plays a significant role in diagnosing most of the cardiac diseases. One cardiac cycle in an ECG signal consists of the P-QRS-T waves. This feature extraction scheme determines the amplitudes and intervals in the ECG signal for subsequent analysis. These features are obtained semi automatically from time-voltage of R, S, T, P, Q features of an Electro Cardiogram signals. The amplitudes and intervals value of P-QRS-T segment determines the functioning of heart of every human. Recently, numerous research and techniques have been developed for analyzing the ECG signal.

We used genetic algorithm for ECG arrhythmia classification. It is used to improve the generalization performance of the SVM classifier. In order to do this, the design of the SVM classifier is optimized by searching for the best value of the parameters that tune its discriminate function, and looking for the best subset of features that optimizes the classification fitness function. This system is automatically select a proper subset of features for optimizing the SVM classifier. We used MAT lab program for this method. After the running program, we get genetic algorithm - support vector method graph for ECG arrhythmia classification.

KEYWORDS: ECG, Arrhythmia, Support Vector Machine, Genetic Algorithm

Original Article

I .INTRODUCTION

Classification of electrocardiograms (ECG) into different disease categories is a complex pattern recognition task. However, the analysis of electrocardiogram signals is the most effective available method for diagnosing cardiac arrhythmias. Computer-based classification of ECGs can provide high accuracy and offer a potential of an affordable cardiac abnormalities mass screening. Successful classification is achieved by finding the characteristic shapes of the ECG that discriminate effectively between the required diagnostic Categories.[1] A typical heart beat is identified from the ECG and the component waves of the QRS, T, and possibly P waves are characterized using measurements such as magnitude and duration. Datasets used for training and test of automated classification of ECG signals include many different features. However, one of the most popular and useful databases is the MIT-BIH. We used this database to various algorithms for arrhythmia detection and classification. Several methods have been proposed for the classification of ECG signals.[4] The arrhythmia of the heart is determined by the shape of the ECG waveform, which contains important information to different types of disease arises in the heart. However, the ECG signals are not regular in nature and it arises randomly at different time

intervals during a day. Sensor data mainly contain noise and distortion and hence before processing these data, the sensor electrocardiogram (ECG) data we need to remove as much noise, distortion present. So proper analyzing the interval between waveforms is needed [2].

Support vector machine and its types is described in section II. In section III, feature extraction and selection is described. Also the dataset is described in detail in this section. In section IV, feature reduction is described. The GA-SVM algorithm is described in section V. Experimental result is given in section VI.

II. SUPPORT VECTOR MACHINE CLASSIFICATION

In this section, a brief description of Support Vector Machines is mentioned. (SVMs) are very popular and powerful in pattern learning because of supporting high dimensional data and, at the same time, providing good generalization properties. SVMs have many usages in pattern recognition. At the beginning, SVM was formulated for two-class (binary) classification problems [6].

A. Binary Support Vector Machine Formulation

Let $X=\{(x_i, y_i)\}^n_{i=1}$ be a set of *n* training samples, where $x_i \in R^m$ is an m-dimensional sample in the input space, and $y_i \in \{-1,1\}$ is the class label of sample x_i. SVM finds the optimal separating hyper plane (OSH) with the minimal classification errors. The linear separation hyper plane is in the form of
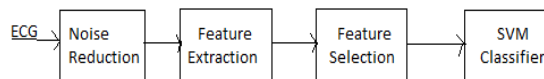
$$f(x) = w^T x + b,$$

Where *w* and *b* are the weight vector and bias, respectively[5].

B. Multi Class Support Vector Machine

SVMs are usually binary classifiers, but, the classification of ECG signals mostly involves more than two classes. So, for this purpose, a number of multiclass classification strategies can be adopted. The most popular ones are the one-against-all (OAA) and the one-against-one (OAO) strategies. We use OAO for ECG multi class classification. The one against one constructs (n (n-1)/2)decision functions for all the combinations of class pairs[6].

III. FEATURE EXTRACTION AND SELECTION

In this section we will explain the characteristics of the extracted feature from the ECG signals and the procedure designed for the extraction. Figure 1, presents the block diagram of the proposed arrhythmia classification. Details of each stage are explained later.

Figure 1: Block Diagram of Proposed Arrhythmia Classification [2]

A. Dataset Description

In our experiment we are taking the ECG data as the basic signal for classification. In recent researches, the annotated ECG records, available at the MIT-BIH arrhythmia database, have been widely used for the evaluation of the

performance of different classifiers. The database has 48 records with each record being an ECG signal for the duration of 30 minutes. Each data was recorded in two channels, modified limb lead II and modified lead VI. The database contains approximately 109,000 beat labels. ECG signals MIT-BIH Database are described by a text header file (.hea), a binary file (.dat) and a binary annotation file (.atr) [6].
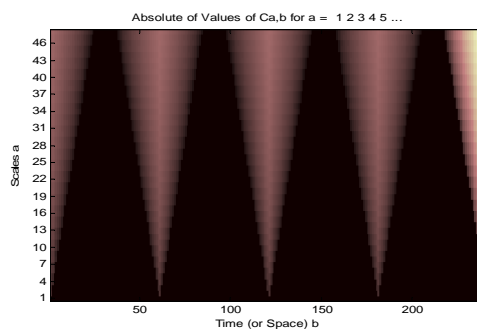
In particular, the considered beat types refer to following classes: normal sinus rhythm (N), right bundle branch block(RB), left bundle branch block (LB), and paced beat (P). The beats were selected from the recording of following patients, which is shown in table 1.

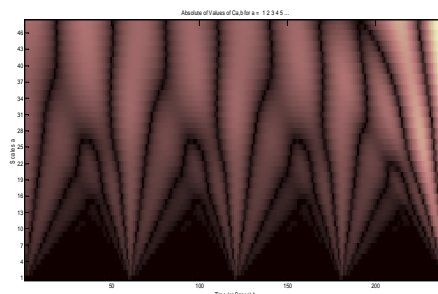**Table 1: Data Set Descriptions and Numbers Used in the Simulation**

Cl. No.	Record Ex.	N	L	R	A	V	F	P	F
1	100	2239			33		1		
2	105	2526				41			
3	107						59		
4	217	244					162	1542	260
5	111		2123			1			
6	212	923		1825					

B. Noise Reduction

In the first stage of feature extraction, a wavelet transformation is performed in order to reduce noises. The wavelet transform allows processing non-stationary signals such as ECG signal. This is possible. Figure shows the ECG signal after reducing noises. We use fifth level of wavelet [12].

Figure 2: Input Signal after Using Wavelet Transform

(First Level)

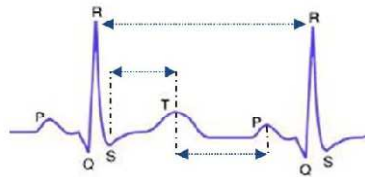**Figure 3: Input Signal after Using Wavelet Transform**

(Fifth Level)

C. Feature Description

For each signal nineteen temporal features such as R-R interval, PQ interval, PR interval, and PT interval and three morphological features are recognized. These features are manually extracted for each beat and put into a separate vector. Each vector is tagged with one the four possible labels N, P, LB, RB [6].

Features have been extracted including the time and voltage of Q/R/S/T/P and time interval for each of 5 features from the next feature such as RS/ ST/ QR as a mentioned in figure 4 and also the difference of voltage in these features such as V (Q)-V(S). The three morphological features by computing the maximum and the minimum values of a beat in ECG signal. Signals of each beat are scaled, using the following formula, such that the range of every signal is between zero and one [12].

$$F(t) = \frac{f(t) - min(t)}{Max(t) - min(t)}$$

The minimum and maximum voltages between the first and the second R feature is computer first and the normalization action is performed [0 1]. As mentioned before, we considered percent that are higher than 0.2, 0.5 and 0.8 as three features. Six of the 22 features, called basic features, are: R1, S, T, P, Q, R2 and the rest are called derived features. The derived features are calculated using the basic features via a semiautomatic procedure.

Figure 4: Sample Features, ST Interval, TP Interval and RR Interval [6]

The features which are used in the simulation are described in table 2.

Table 2: Features Descriptions Used in the Simulation

Feature No.	Description
01	X(R1)
02	V(R1)
03	X(S)
04	V(S)
05	X(T)
06	V(T)
07	X(P)
08	V(P)
09	X(Q)
10	V(Q)
11	X(R2)
12	V (R2)
13	X(R2) – X(R1)
14	V(R2) – V(R1)
15	X(S) – X(R1)
16	X(T) – X(S)
17	X(P) – X (T)
18	X(Q) – X(P)
19	X(R2) – X(Q)

IV. FEATURE REDUCTION

Many researchers said that not all the features are useful for classification. Some features may act as noises and, hence, reduce the classification accuracy. In this paper, feature reduction approaches are adopted. The studied showed that a meta-heuristic-based approach has better performance for classification of ECG arrhythmia than statistical method.

A. Genetic Algorithm

Genetic Algorithm is one prospective option for feature reduction. It can solve every optimization problem which can be described with the chromosome encoding. It solves problems with multiple solutions. The details of our implementation of GA are described as follows:

Algorithm: Genetic Algorithm

Input: Training Data

Output: Useful Features

Step0: initialize parameters (e.g. population size, crossover Rate, mutation rate and the maximum number of population

generation.)

Step1: create initial population randomly (*P (0)*).

Step2: evaluate current population (compute fitness of all

chromosomes).

Step3: while (termination condition not satisfied) do [step

4-8]

Step4: select *P(t)* from *P(t+1)* [perform selection]

Step5: recombine *P (t)* [perform mutation and crossover]

Step6: evaluate current population (compute fitness of all

Chromosomes).

Step7: $t = t + 1$

Step8: go to Step 3

V. THE PROPOSED GENETIC – SVM

In this section, the proposed Wavelet based Genetic-SVM system for the classification of ECG signals is described. The aim of this system is to reduce noise from incoming signal and automatically select a proper subset of features for optimizing the SVM classifier. The upcoming figure shows the procedure followed.

A. Genetic Set Up:

The procedure of SVM classification system is as follows:

Step 1) Randomly generates an initial population of size 50.

Step 2) For each chromosomes of the population, train P(0) SVM Classifiers.

Step 3) Using OAO (multi-class SVM), compute the fitness of each chromosome (subset of features).

Step 4) Directly Select some individuals from the current population based on the fitness values and regenerate new individuals from old ones.

Step 5) If the maximum number of iteration is not yet reached, return to step 2.

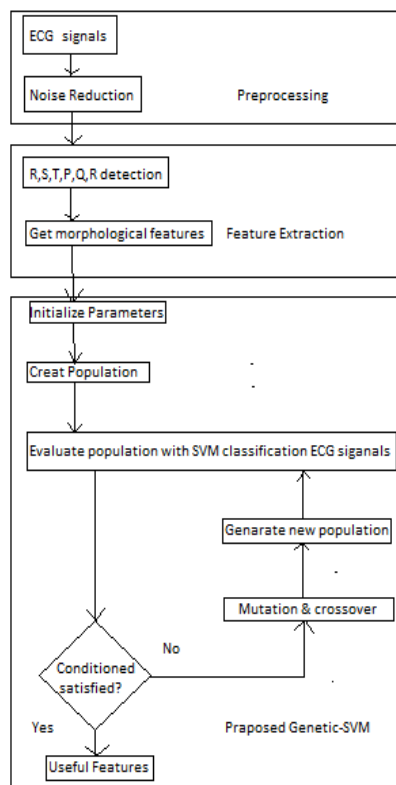**Steps 6)** Select the chromosome with the best fitness value as the desired subset of features.

Step 7) Classify the ECG Signals with the trained SVM

Sensitivity refers to the rate of correctly classified positive. Sensitivity may be referred as a True Positive Rate. Sensitivity should be high for a classifier

Sensitivity = TP / (TP+FN)

True Positive (TP): arrhythmia detection coincides with decision of physician

False Negative (FN): system labels an arrhythmia as healthy

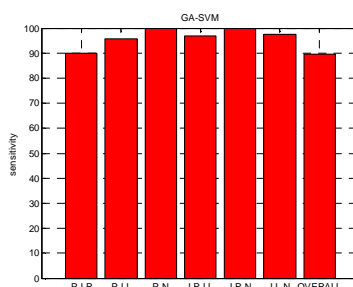**Figure 5: The Wavelet Based Genetic –SVM Approach [2]**

VI. EXPERIMENTAL RESULT

For the evaluation of the proposed method, 50% of all the data of the MIT-BIH dataset are used for training the composed system and the rest are used for the evolution. The genetic-SVM classifier is used to the best subset of features

that can optimize the SVM classifier. Wavelet transform reduces noise signal and helps to improve result. The results of this experiment are presented in the Table 3, 4 and 6. The graph of different approaches is mentioned bellow figure 6, 7 and 8. The comparison of parameters of all methods is shown in table 6.

Table 3: The Arrhythmia Classification Result using GA-SVM Method
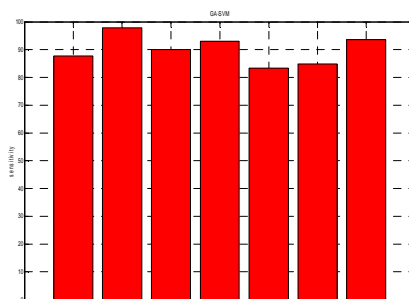
P,LR	P,LL	P,N	LR,LL	LR,N	LL,N	Overall
89.90	95.11	100	95.56	100	98.01	91.13

Figure 6: Result of GA- SVM Method

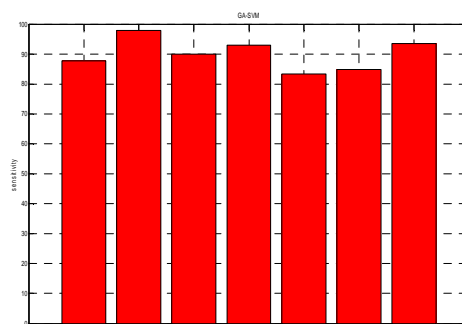**Table 4: The Arrhythmia Classification Result Using GA-SVM Method**

P,LR	P,LL	P,N	LR,LL	LR,N	LL,N	Overall
89.90	98.11	90.01	93.11	84.01	85.01	93.13

Figure 7: Result of Wavelet Based GA-SVM Method (First Level)

Table 5: The Arrhythmia Classification Result Using GA-SVM Method

P,LR	P,LL	P,N	LR,LL	LR,N	LL,N	Overall
89.90	99	90	92.11	84.01	85.01	95.12

Figure 8: Result of Wavelet Based GA-SVM Method (Fifth Level)

Table 6: Comparisons of Parameters of Different Methods

Parameter	SVM	SVM-GA	Wavelet-SVM	Wavelet-SVM-GA
Overall sensitivity	0.8812	0.9113	0.8913	0.9512
SVM sensitivity	0.8805	0.9201	0.9311	0.9302
Executing time	7.5s	35.14s	40.1s	42.12s
Efficiency	0.7502	0.7812	0.8012	0.9021
Quality factor	0.92	0.93	0.94	0.96

VII. CONCLUSIONS AND FUTURE WORK

In this paper, a new method based on support vector machines for feature selection of electrocardiogram signals is proposed. Experimental results show that feature selection greatly improves the quality of classification. This is because some features may act as noises and negatively affect the outcome. A flexible and effective GA-SVM method for classification of ECG signal was presented and showed to be effective on MIT-BIH dataset to recognize arrhythmia classes. For future work, diseases of ECG can be identified.

ACKNOWLEDGEMENTS

REFERENCES

1. *P. Keerthi Priya1, Dr. G. Umamaheswara Reddy, ``MATLAB Based GUI for Arrhythmia Detection Using Wavelet Transform'', International Journal of Advance Research in Electrical, Electronics and International Engineering, An ISO 3297, 2007 certified organization, vol.4, Issue 2, February 2015.*

2. *Prof. Pratik Tawde, Prof. Arun Chavan, ``Use Of Genetic-SVM For ECG Arrhythmia Classification'', International Journal of Advance Research In Science And Engineering, IJARSE, Vol. No.4, Issue 04, April 2015.*

3. *Yutana Jewajinda, Prabhas Chongstitvatana, ``A Parallel Genetic Algorithm For Adaptive Hardwaremand Its Application To ECG Signal Classification'', Neural Comput and Applic (2013) 22:1609–1626, Received: 21 April 2011 / Accepted: 21 May 2012 / Published online: 12 June 2012.*

4. *V.Rathikarani, P.Dhanalakshmi, ``Automatic Classification of ECG signal for Identifying Arrhythmia'', International Journal of Advance Research In Science And Engineering, Volume 3, Issue 9, September 2013.*

5. *Ritaban Kirtania, Kalyani Mali, `` Cardiac Arrhythmia Classification using Optimal Feature Selection and K-Nearest Neighbour Classifier '', International Journal of Advance Research In Science And Engineering, Volume 5, Issue 1, January 2015.*

6. *Khazaee, A. Ebrahimzadeh, ``Classification Of Electrocardiogram Signals With Support Vector Machines and Genetic Algorithm Using Power Spectral Features '', Biomedical Signal Processing and control, Science Direct, Vol. 5, Issue 4, October 2010, Pages 252-263.*

7. *S. Z. Mahmoodabadi, A. Ahmadian,and M. D. Abolhasani, ``ECG Feature Extraction using Daubechies Wavelets'', Proceedings of the fifth IASTED International Conference on Visualization, Imaging and Image Processing, pp. 343-348, 2005.*

8. *Juan Pablo Martínez, Rute Almeida, and Pablo Laguna, ``A Wavelet Based ECG Delineator: Evaluation on Standard Databases'', IEEE Transactions on Biomedical Engineering, Vol. 51, No. 4, pp. 570-581, 2004.*

9. Krishna Prasad and J. S. Sahambi, ``Classification of ECG Arrhythmias using Multi-Resolution Analysis and Neural Networks'', IEEE Transactions on Biomedical Engineering, vol. 1, pp. 227-231, 2003.

10. Cuiwei Li, Chongxun Zheng and Changfeng Tai, ``Detection of ECG Characteristic Points using Wavelet Transforms'', IEEE Transactions on Biomedical Engineering, Vol. 42, No. 1, pp. 21-28, 1995.

11. C. Saritha, V. Sukanya and Y. Narasimha Murthy, ``ECG Signal Analysis Using Wavelet Transforms'', Bulgarian Journal of Physics, vol. 35, pp. 68-77, 2008.

12. Qibin Zhao and Liqing Zhan, ``ECG Feature Extraction and Classification Using Wavelet Transform and Support Vector Machines'', International Conference on Neural Networks and Brain, ICNN and B '05, vol. 2, pp. 1089-1092, 2005.

13. T. Inan, L. Giovangrandi, and J. T. A. Kovacs, "Robust neural network based classification of premature ventricular contractions using wavelet transform and timing interval features," IEEE Trans. Biomed. Eng., vol.53, no. 12, pp. 2507–2515, Dec. 2006.

14. Foody, M.G., Mathur, A., "A Relative Evaluation of Multiclass

15. Image Classification by Support Vector Machines", IEEE Transactions on Geoscience and Remote Sensing, 42, 1335– 1343, 2004.

16. [iaomin Xu and Ying Liu, ``ECG QRS Complex Detection Using Slope Vector Waveform (SVW) Algorithm'', Proceedings of the 26th Annual International Conference of the IEEE EMBS, pp. 3597-3600, 2004.

17. A. Gharaviri, F. Dehghan, M. Teshnelab, H. Abrishami, ``comparision of neural network anfis and SVM classification for PVC arrhythmia detection'', Proceedings of the Seventh International Conference on Machine Learning and Cybernetics, Kunming, 12-15 July 2008.

18. Jala A. Nasari, Mahmoud Naghibzadeh, H. Sadoghi Yazdi, ``ECG Arrhythmia Classification using Support Vector Machines and Genetic Algorithm'', Proceeding of the third UK Sim European Symposium on Computer Modeling and Simulation, 2009.

19. Christopher J.C. Burgrs, ``A Tutorial on Support Vector Machines for Pattern

20. Recognition'', Proceeding of the Bell Laboratories, Lucent Technologies, Kluwer Academic Publishers, Boston. Manufactured in The Netherlands, 2010.

21. B. Castro, D. Kogan and A. B. Geva, ``ECG feature extraction using optimal mother wavelet'', The 21st IEEE Convention of the Electrical and Electronic Engineers in Israel, pp. 346-350, 2000.

22. Anuja sas, Shreetam behera, "Identification of Tachycardia and Bradycardia

23. Heart Disorders using Wavelet Transform based QRS Detection" ,

24. International Journal of Image Processing and Vision Sciences,Volume-2 Issue-1, pp.2278–1110 , 2013.